ISSN 2788-7995 (Print)
ISSN 3006-0524 (Online)

ШӘКӘРІМ УНИВЕРСИТЕТІНІҢ ХАБАРШЫСЫ ТЕХНИКАЛЫҚ ҒЫЛЫМДАР

ҒЫЛЫМИ ЖУРНАЛ

ВЕСТНИК УНИВЕРСИТЕТА ШАКАРИМА ТЕХНИЧЕСКИЕ НАУКИ

НАУЧНЫЙ ЖУРНАЛ

BULLETIN OF SHAKARIM UNIVERSITY TECHNICAL SCIENCES

SCIENTIFIC JOURNAL

№3 (19)
2025

https://vestnik.tech.shakarim.kz

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ ҒЫЛЫМ ЖӘНЕ ЖОҒАРЫ БІЛІМ МИНИСТРЛІГІ
МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РЕСПУБЛИКИ КАЗАХСТАН

SHAKARIM UNIVERSITY
1934

ШӘКӘРІМ УНИВЕРСИТЕТІНІҢ
Х А Б А Р Ш Ы С Ы
ТЕХНИКА ҒЫЛЫМДАР

ҒЫЛЫМИ ЖУРНАЛ

В Е С Т Н И К
УНИВЕРСИТЕТА ШАКАРИМА
ТЕХНИЧЕСКИЕ НАУКИ

НАУЧНЫЙ ЖУРНАЛ

B U L L E T I N
OF SHAKARIM UNIVERSITY
TECHNICAL SCIENCES

SCIENTIFIC JOURNAL

№ 3 (19) 2025

Семей, 2025

Ғылыми журнал
«Шәкәрім Университетінің Хабаршысы. Техникалық ғылымдар сериясы»

№ 3 (19) 2025

Меншік иесі:
«Шәкәрім университеті» Коммерциялық емес акционерлік қоғамы

1997 жылдан бастап шығарылады
Кезеңділігі: тоқсан сайын (жылына 4 рет)

Журнал Қазақстан Республикасы Ақпарат және қоғамдық даму министрлігінің
Ақпарат комитетінде тіркелген
Есепке қою туралы куәлік № KZ93VPY00033663 19.03.2021 ж.

Жазылу индексі: 76172

Редакция құрамы:
Евлампиева Е.П. – редактор
Семейская З.Т. – редактор

Редакцияның мекен-жайы:
071412, Абай облысы, Семей қ., Глинки к-сі, 20а, каб. 506
Байланыс телефоны: 8(7222)31-32-49
Электрондық пошта: rio@shakarim.kz

Научный журнал
«Вестник Университета Шакарима. Серия технические науки»

№ 3 (19) 2025

Собственник:
Некоммерческое акционерное общество «Шәкәрім университет»

Издается с 1997 года
Периодичность: ежеквартально (4 раза в год)

Журнал зарегистрирован в Комитете информации Министерства информации
и общественного развития Республики Казахстан
Свидетельство о постановке на учет № KZ93VPY00033663 от 19.03.2021 г.

Подписной индекс: 76172

Технические редакторы:
Евлампиева Е.П.
Семейская З.Т.

Адрес редакции:
071412, область Абай, г. Семей, ул. Глинки, 20А, каб. 506
Контакты: телефон: 8(7222)31-32-49
Электронная почта: rio@ shakarim.kz

МАЗМҰНЫ – СОДЕРЖАНИЕ

АВТОМАТТАНДЫРУ ЖӘНЕ АҚПАРАТТЫҚ ТЕХНОЛОГИЯЛАР

АВТОМАТИЗАЦИЯ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

ТЕХНИКАЛЫҚ ФИЗИКА ЖӘНЕ ЖЫЛУЭНЕРГЕТИКАСЫ

ТЕХНИЧЕСКАЯ ФИЗИКА И ТЕПЛОЭНЕРГЕТИКА

ХИМИЯЛЫҚ ТЕХНОЛОГИЯ

ХИМИЧЕСКАЯ ТЕХНОЛОГИЯ

Басуға жіберілген күні 30.09.2025 ж. Пішімі 60х84 1/8
Шартты баспа табағы 41,6
Таралымы 100 дана. Бағасы келісімді.

Техникалық редакторы: Е.П. Евлампиева, З.Т. Семейская
Безендіруші: С.Т. Мырзабеков

Журнал Қазақстан Республикасы Ақпарат және қоғамдық даму министрлігінің
Ақпарат комитетінде тіркелген
Есепке қою туралы куәлік № KZ93VPY00033663 19.03.2021 ж.

Жылына 4 рет шығады

Құрылтайшысы: «Шәкәрім университеті»
коммерциялық емес акционерлік қоғам

Шәкәрім университетінің
баспаханасында басылды

Редакцияның мекен-жайы: 071412, Абай облысы,
Семей қаласы, ул. Глинки, 20 А
Тел.: +7 (7222) 31-32-49, эл.почта: rio@ shakarim.kz

Жантасова Женискүл Зейнешовна – компьютерлік модельдеу және ақпараттық технологиялар кафедрасының доценті; С. Аманжолов атындағы Шығыс Қазақстан университеті; e-mail: Zheniskul_z@mail.ru.

Information about the authors

Andrey Bondarovich – candidate of Geographical Sciences; Associate Professor of the Department of economic geography and cartography, Altai State University; e-mail: a9130262571@gmail.com. ORCID: https://orcid.org/0000-0001-7161-2686.

Almasbek Maulit* – doctoral student of Shakarim University; head of the Technopark «Shygys Bastau», EKU. S. Amanzholova; e-mail: maulit.almas@gmail.com. ORCID: https://orcid.org/0000-0002-0519-3222.

Igor Alexandrovich Ocheredko – researcher of the Technopark «Shygys Bastau», East Kazakhstan University. S. Amanzholova; e-mail: egor007kz@mail.ru.

Zheniskul Zeyneshovna Zhantasova – associate professor of the Department of computer modeling and information technologies; East Kazakhstan University. S. Amanzholova; e-mail: Zheniskul_z@mail.ru.

Поступила в редакцию 07.06.2025
Поступила после доработки 11.08.2025
Принята к публикации 12.08.2025

https://doi.org/10.53360/2788-7995-2025-3(19)-13

МРНТИ: 50.41.25

А.М. Искакова[1*], О.С. Салыкова[1], Н.А. Дидарбекова[2], А.А. Артыкбаева[1]
[1]Костанайский региональный университет имени Ахмет Байтұрсынұлы,
110000, Республика Казахстан, г. Костанай, ул.А.Байтурсынова, 47
[2]Национальный центр тестирования
Министерства науки и высшего образования Республики Казахстан
010011, Республика Казахстан, г. Астана, улица Родниковая 1/1
*e-mail: n.a.almira.24@gmail.com

ШКАЛИРОВАНИЕ ТЕСТОВЫХ ЗАДАНИЙ ПО ФИЗИКЕ ДЛЯ КОМПЬЮТЕРНОГО АДАПТИВНОГО ТЕСТИРОВАНИЯ НА ОСНОВЕ МОДЕЛИ РАША

Аннотация: *Адаптивное тестирование представляет собой один из наиболее эффективных подходов к цифровой оценке знаний, обеспечивая индивидуализацию процесса за счёт автоматического подбора заданий в зависимости от уровня подготовленности обучающегося. Основными элементами для реализации такого тестирования являются: банк шкалированных заданий, алгоритм адаптации и специализированное программное обеспечение. Формирование качественного банка заданий требует предварительного психометрического анализа, позволяющего оценить их пригодность к использованию в адаптивных системах.*

Настоящая статья посвящена эмпирическому анализу набора тестовых заданий по физике с использованием модели Раша. В ходе исследования проведена апробация заданий на репрезентативной выборке учащихся и выполнено шкалирование с применением программы Winsteps. Для каждого задания определены параметры трудности, показатели согласия с моделью и коэффициенты корреляции. Выявлены задания, не соответствующие требованиям адаптивного тестирования ‒ они были исключены из итогового банка. В результате сформирован набор заданий, обладающих устойчивыми статистическими характеристиками и пригодных для дальнейшего использования в компьютерных адаптивных системах оценки знаний.

Представленные результаты подтверждают возможность интеграции полученного банка заданий в информационные образовательные системы и цифровые платформы. В дальнейших публикациях планируется представить алгоритмы построения адаптивного тестирования в реальном времени, а также разработку программного обеспечения для автоматической генерации тестов с учётом шкалированных параметров. Проведённая работа служит основой для создания эффективных цифровых инструментов оценки учебных достижений.

Ключевые слова: *адаптивное тестирование, модель Раша, психометрический анализ, шкалирование заданий, информационные технологии в образовании, Winsteps, автоматизированная оценка знаний.*

Введение

Компьютерное адаптивное тестирование (КАТ) является одним из наиболее перспективных направлений цифровизации оценки знаний. В отличие от традиционных методов, КАТ позволяет автоматически подбирать задания индивидуально под уровень подготовленности каждого испытуемого, тем самым обеспечивая точность измерений и сокращение времени тестирования [1]. В педагогических измерениях особое место занимает модель Раша, обеспечивающая шкалирование заданий и испытуемых на единой шкале, независимо от конкретного содержания теста. Использование данной модели в сочетании с современным программным обеспечением, таким как Winsteps, открывает возможности для формирования надёжных и объективных банков заданий. Одной из ключевых задач, стоящих перед исследователями в области КАТ, является обеспечение качества тестовых заданий, в частности: их соответствие по уровню трудности, согласованность статистик согласия с моделью, функционирование дистракторов. На практике это требует многоступенчатого психометрического анализа и статистической верификации полученных результатов [2, 3].

Целью настоящего исследования является проведение психометрического анализа тестовых заданий по физике с использованием модели Раша для формирования банка шкалированных заданий, пригодных к внедрению в цифровые адаптивные системы тестирования.

Для достижения поставленной цели необходимо решить следующие задачи: провести апробацию тестовых заданий на репрезентативной выборке, выполнить анализ параметров трудности и статистик согласия с применением модели Раша, отобрать задания, удовлетворяющие критериям, обосновать применение полученных результатов в системах цифровой оценки знаний.

Материалы и методы

В педагогических измерениях анализ тестовых заданий основывается на данных, полученных эмпирическим путем. Психометрический анализ тестовых заданий - это процесс оценки статистических характеристик заданий теста с целью определения их качества и соответствия целевому назначению теста [4]. Для получения точных и объективных измерений латентных переменных в данном исследовании авторы использовали программу Winsteps. На рисунке 1 представлены этапы процесса формирования банка шкалированных тестовых заданий с последующей его интеграцией в систему компьютерного адаптивного тестирования (КАТ) [5].

Рисунок 1 – Этапы формирования банка заданий для КАТ.

Первым шагом процесса исследования теста является создание и редактирование матриц с результатами тестирования. Поскольку любое педагогическое измерение требуют обязательной корректировки исходных матриц, из полученных данных были исключены результаты экстремальных испытуемых и сформированы редуцированные матрицы для обработки в Winsteps.

В теории педагогических измерений рассматривают несколько мер трудности заданий. Первую меру трудности заданий можно рассмотреть, как отношение числа неправильных ответов на задание к числу испытуемых.

$$q_j=W_j/N,$$

где Wj – число неправильных ответов на задание под номером j и N – число испытуемых. Вторая мера определяется как отношение q_j/p_j, где p_j- доля правильных ответов. Третья мера трудности вычисляется как значение натурального логарифма отношения q_j/p_j. Полученное значение далее корректируется для построения единой шкалы уровня трудности заданий и уровня подготовленности испытуемых. Это и есть процесс шкалирования, проводимый с помощью статистического пакета Winsteps. Скорректированные в процессе шкалирования значения $\ln q_j/p_j$ являются четвертой мерой трудности. Данная мера применяется в качестве окончательной меры трудности и называется параметром трудности задания [6].

Эффективные дистракторы в тестовых заданиях, предназначенных для адаптивного тестирования, должны обладать определенными характеристиками: быть правдоподобными для испытуемых с низким уровнем подготовленности, повышать вероятность выбора неправильного ответа и при этом не снижать дифференцирующую способность задания, не вводить в заблуждение сильных испытуемых. Дистракторный анализ позволяет оценить соответствие дистракторов указанным критериям и является важным этапом процесса формирования базы тестовых заданий [7, 8].

Результаты и обсуждение

В данной работе для проведения апробации было использовано 10 вариантов тестов по физике на основе единой технологической матрицы. Содержание тестовых заданий соответствует программе курса физики для средней общеобразовательной школы. Количество заданий в одном варианте – 30 (всего использовано 300 тестовых заданий). Общее количество участников апробации – 3190. По результатам апробации в двух вариантах тестов были выявлены экстремальные испытуемые. К экстремальным испытуемым относят учащихся ответивших верно на все задания теста, уровень их подготовленности выше уровня трудности проектируемых тестов, соответственно результаты данных учащихся были исключены из исходных матриц.

В данном исследовании интерпретация результатов тестирования будет проводиться на основе модели Раша, в рамках которой, оценка уровня подготовленности испытуемых не зависит от конкретных заданий, используемых в тесте [9]. Аналогично, оценка трудности задания не зависит от конкретных испытуемых, проходящих тест. Анализ заданий будет представлен на примере одного варианта теста.

В таблицах 1 и 2 приведены результаты апробации, а именно общая статистика по испытуемым и тестовым заданиям.

Таблица 1 – Результаты апробации (испытуемые)

	Исходный балл (0-30)	Оценка уровня подготов. (в логит.)	Ошибка измерения уровня подготов.	Статистика согласия			
				$U^{(1)}$	$t^{(1)}$	$U^{(2)}_{(взвеш)}$	$t^{(2)}_{(взвеш)}$
Ср. значение	11,4	-0,55	0,42	1,00	-0,1	1,02	0,0
Ср.квадр.отклон	4,8	0,80	0,05	0,11	0,9	0,18	0,9
Макс.значение	28,0	2,75	0,74	1,37	2,4	2,04	2,4
Мин.значение	2,0	-2,74	0,38	0,74	-2,9	0,62	-2,8

Таблица 2 – Результаты апробации (задания)

	Процент решенных заданий, %	Оценка трудности заданий (в логит.)	Ошибка измерения трудности заданий.	Статистика согласия			
				$U^{(1)}$	$t^{(1)}$	$U^{(2)}_{(взвеш)}$	$t^{(2)}_{(взвеш)}$
Ср. значение	37,9	0,00	0,12	1,00	-0,4	1,02	-0,2
Ср.квадр.отклон	10,6	0,53	0,01	0,10	2,0	0,15	1,9
Макс.значение	61,5	0,92	0,14	1,26	3,4	1,44	3,7
Мин.значение	20,9	-1,12	0,12	0,86	-4,2	0,82	-3,9

Согласно данным представленным в таблицах 1 и 2, можно отметить, что уровень подготовленности учащихся, выполнявших данный вариант оказался невысоким. Средний балл испытуемых, указанный в первом столбце таблицы 1, составляет 11.4 Во втором столбце указанных таблиц представлены значения оценок трудности заданий и уровней подготовленности испытуемых в логитах. Уровень подготовленности учащихся, проходивших апробацию данного теста, значительно отличается: самый подготовленный испытуемый имеет оценку уровня подготовленности Ө=2,75, слабо подготовленный Ө=-2,74. Числовые значения уровней трудности заданий во втором столбце таблицы 2 показывают их *относительную трудность*, так самое легкое задание имеет уровень трудности ß=-1,12, самое трудное ß=0,92. В третьем столбце таблиц содержатся значения среднеквадратичных ошибок вычисления оценок. Из представленных данных следует отметить, что ошибка измерения для заданий невелика, среднее значение ошибки измерения по тестовым заданиям составляет 0,12, по испытуемым 0,42. Высокая точность при оценивании параметров заданий может объяснятся тем, что количество участников тестирования в 10 раз больше количества заданий, а точность результатов обеспечивается достаточным объемом выборки. В последних столбцах таблиц представлены данные о статистиках согласия, которые характеризуют согласие экспериментальных данных, полученных во время апробации с используемой моделью измерения. Как правило - это общие статистики согласия и их стандартизированные версии, математические ожидания значений общих статистик согласия равны 1, стандартизированных – 0. Значения средних статистик согласия представленные в таблицах 1 и 2 указывают на то, что большая часть заданий и испытуемых находятся в хорошем согласии с моделью, о чем свидетельствуют их близкие значения к ожидаемым. Вместе с тем максимальные значения статистик согласия, превышающие критические значения демонстрируют наличие заданий и испытуемых, находящихся в плохом согласии с моделью. Результаты данных заданий будут проанализированы ниже.

Статистические данные по заданиям теста приведены в таблице 3. Задания расположены в порядке убывания уровня трудности. В первом столбце указан номер задания, в следующих двух столбцах оценка трудности задания в логитах и ошибка измерения трудности задания. Согласно представленным данным самые трудные задания - №19 и №29: их уровни трудности ß=0,92 и 0,87 соответственно, ошибка измерения у обоих заданий равна – 0,14. Самое легкое задание №12, его уровень трудности ß=-1,12, ошибка измерения – 0,12. Далее в 4-х столбцах представлены значения статистик согласия U(1), t(1), U(2), t(2), их средние значения представлены в таблицах 1 и 2. Коэффициент корреляции баллов по заданию с общим баллом по тесту приведен в последнем столбце.

Таблица 3 – Результаты апробации

Номер задания	Трудность задания (логит)	Ошибка измерения	Статистики согласия				Коэф. корреляции
			$U^{(1)}$	$t^{(1)}$	$U^{(2)}$	$t^{(2)}$	
19	0.92	0,14	1,05	0,7	1,03	0,3	0,26
29	0.87	0,14	1,26	3,2	1,44	3,7	-0,1
1	0.66	0,13	1,03	0,5	0,99	-0,1	0,3
24	0.62	0,13	1,24	3,4	1,36	3,6	0,04
25	0.59	0,13	1,14	2,1	1,33	3,4	0,14
28	0.55	0,13	1,05	0,8	1,03	0,4	0,28
11	0.43	0,13	1,03	0,6	1,04	0,5	0,30
21	0.41	0,13	1,06	1,0	1,07	0,9	0,27
22	0.38	0,13	1,15	2,6	1,20	2,5	0,15
23	0.37	0,13	1,07	1,3	1,10	1,3	0,25
30	0.36	0,13	0,99	-0,2	1,00	0,0	0,35
17	0.35	0,12	1,05	0,9	1,11	1,4	0,27
20	0.14	0,12	1,08	1,6	1,10	1,5	0,25
15	0.13	0,12	0,91	-1,9	0,89	-1,8	0,45
6	0.02	0,12	0,97	-0,7	0,98	-0,2	0,38
26	-0.01	0,12	0,96	-1,0	0,95	-0,9	0,40
5	-0.04	0,12	0,92	-1,9	0,92	-1,5	0,44
13	-0.15	0,12	0,97	-0,8	0,95	-1,0	0,39
27	-0.17	0,12	1,03	0,7	1,03	0,5	0,32
18	-0.30	0,12	0,99	-0,3	0,98	-0,4	0,36
7	-0.37	0,12	0,86	-4,0	0,84	-3,4	0,51
4	-0.43	0,12	0,89	-3,2	0,87	-2,7	0,48
3	-0.44	0,12	0,90	-2,8	0,94	-1,3	0,45
9	-0.46	0,12	0,91	-2,6	0,90	-2,2	0,45
16	-0.50	0,12	0,96	-1,0	0,97	-0,7	0,39
8	-0.57	0,12	0,88	-3,4	0,86	-3,0	0,48
2	-0.68	0,12	0,86	-4,2	0,82	-3,9	0,52
14	-0.72	0,12	0,92	-2,3	0,89	-2,2	0,44
10	-0.84	0,12	1,01	0,2	0,99	-0,1	0,33
12	-1.12	0,12	0,95	-1,2	0,99	-0,2	0,38

Анализ данных таблицы 3 показывает, что эмпирическое распределение трудностей заданий, полученное в ходе апробации, отличается от предполагаемого (теоретического) распределения. Вероятной причиной указанного несоответствия является многомерность теста, обусловленная содержательной разнородностью заданий по физике. Для выяснения причин расхождения эмпирических и теоретических данных необходимо провести углублённый анализ целой группы заданий и пересмотреть их распределение по уровням трудности.

Для выявления некорректных тестовых заданий могут быть использованы статистики согласия, позволяющие оценить соответствие ответов испытуемых ожиданиям модели. В литературе по педагогическим измерениям рекомендуют для общих статистик $U^{(1)}$ и $U^{(2)}$ выбирать в качестве правого критического значения – 1,5, левого – 0,5, а для стандартизированных статистик $t^{(1)}$ и $t^{(2)}$ – значения -2 и 2 [10]. Данные таблицы 3 указывают на то, что большинство заданий теста имеют хорошее согласие с моделью. Вместе с тем, тест содержит задания, параметры которых значительно превышают рекомендованные, например тестовое задание №29. Значения статистик согласия $t^{(1)}$ и $t^{(2)}$ у данного задания больше правых критических, следует обратить внимание и на отрицательное значение коэффициента корреляции. Аналогичные проблемы с завышенными значениями статистик согласия наблюдаются у заданий №22,24,25. Из перечисленных тестовых заданий наиболее проблемным является задание №29. Низкие результаты хорошо подготовленных испытуемых и отрицательный коэффициент корреляции, могут свидетельствовать о наличии проблемы в условии задания или коде правильного ответа. Анализ тестового задания №29, включая дистракторный анализ будет представлен ниже.

Из рисунка 2 видно, что график задания № 29 сильно отличается от теоретической кривой: большинство эмпирических точек находятся вне модельной кривой и часть из них вне доверительного коридора. Более того, испытуемые с низким уровнем подготовленности выполнили это задание лучше, чем испытуемые с высоким уровнем подготовленности. Задание абсолютно не дифференцирует испытуемых по уровню подготовленности. Это задание находится в плохом согласии с ответами испытуемых на другие задания теста. На рисунке 2 для сравнения приведен график задания №13. Данное задание является заданием среднего уровня трудности, все критерии согласия которого находятся в допустимых пределах. Уровень трудности задания – (-0,15). Коэффициент корреляции составил 0,39, что говорит о хорошей дифференцирующей способности задания. Задание №13 демонстрирует хорошее согласие с моделью.

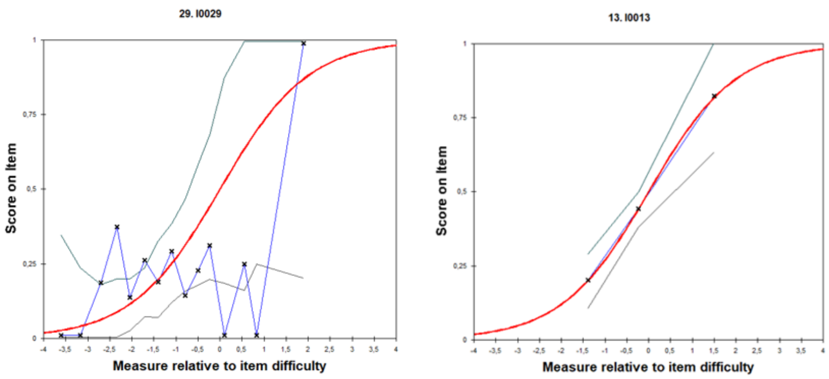

Рисунок 2 – Графики задании № 29 и № 13

Как отмечалось ранее, необходимым условием для определения качества тестовых заданий является проведение дистракторного анализа. Дистракторный анализ позволяет определить эффективность подбора всех ответов, прилагаемых к каждому заданию теста. Без дистракторного анализа ни тестов, ни педагогических измерений не бывает [8]. По итогам дистракторного анализа было установлено, что большинство дистракторов теста сформулированы корректно и функционируют должным образом. Согласно требованиям тестологии, дистрактор считается корректным и работающим если его выбирают не менее 5% учащихся. На рисунке 3 приведены данные по вариантам ответов на задания № 13 и № 29.

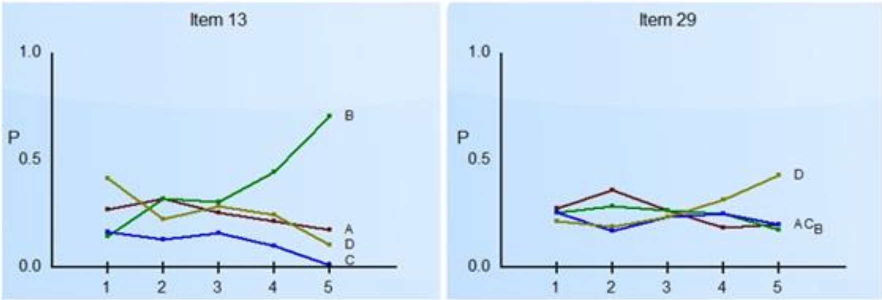

Рисунок 3 – Выбор вариантов ответов

В левой части рисунка 3 показана графическая интерпретация дистракторов для задания № 13. Зеленая линия (B) указывает на правильный ответ. 140 учащихся дали верный ответ на данное задание, что составляет 48% от общего числа испытуемых. Как видно из представленного рисунка сильные учащиеся правильно ответили на тестовое задание: чем выше уровень подготовленности испытуемых, тем выше у них вероятность правильного ответа. Слабые учащиеся выбрали неверные варианты ответов (A,C,D). Еще одним подтверждением верного функционирования вариантов ответов является отрицательное значение коэффициента корреляции у дистракторов, и его положительное значение у правильного ответа.

По результатам дистракторного анализа задания № 29 можно отметить, что правильный вариант ответа (вариант C) выбрали 75 испытуемых, что составляет лишь 22% от общего количества, отрицательное значение коэффициента корреляции правильного ответа свидетельствует о том, что выбирали правильный ответ учащиеся с низким уровнем подготовленности, а учащиеся с высоким уровнем подготовленности выбирали вариант D. Для выяснения причин такого функционирования задания рассмотрим данное задание с его подробным решением. Задание № 29 относится к заданиям на основе контекста. В качестве контеста здесь представлено описание лазерной указки с параметрами. В таблице 4 приведен контекст, условие задания и ход решения задания с пояснениями.

Таблица 4 – Тестовое задание № 29 с решением

Контекст	Задание №29	Решение задания
Лазерная указка представляет собой компактное устройство, создающее узкий световой пучок за счёт вынужденного излучения. Часто она используется для демонстраций и указаний на объекты на расстоянии. Лазеры могут генерировать свет в широком диапазоне длин волн. В стандартной лазерной указке мощность светового излучения составляет 5 мВт, а длина волны – 650 нм (красный лазер). Для питания устройства используется батарея с ЭДС 1,5 В. При работе указка потребляет электрическую мощность 0,1 Вт.	29. При непрерывной работе лазерной указки в течение 3 минут выделится тепловая энергии в количестве A) 0,9 Дж B) 16,2 Дж C) 17,1 Дж D) 18,0 Дж Правильный ответ: C	Находим общую энергию, потребленную указкой: Энергия (E) = Мощность (P) × Время (t) E = Pпотр × t = 0,1 Вт × 180 с = **18 Дж** Находим энергию излучения лазера: Eизл = Pизл × t = 0,005 Вт × 180 с = 0,9 Дж Находим тепловую энергию: По закону сохранения энергии, потребленная энергия идет на излучение и тепло: E = Eизл + Q Q = E - Eизл = 18 Дж - 0,9 Дж = **17,1** Дж

Повторная экспертиза тестового задания № 29 показала, что задание корректное, код правильного ответа соответствует варианту C. Вероятной причиной низких характеристик задания является сложность представленной темы, невнимательность испытуемых при его решении. Как представлено в таблице 4, вариант D, является хорошо подобранным правдоподобным, но неправильным ответом, и представляет собой только общую потребляемую энергию. Для расчета тепловой энергии следует применить закон сохранения энергии, который не применили подготовленные учащиеся. Таким образом, в данном задании следует отметить мастерство разработчика заданий, который смог хорошо подобрать неправильные, но правдоподобные варианты ответов. Следует также отметить, что задание №29 – это задание высокого уровня трудности. Как правило задания высокого уровня трудности характеризуют воспроизведение более сложных знаний и навыков, распознавание более сложных моделей заданий, интегрирование знаний, умений и навыков.

В рамках данного исследования программой Winsteps была рассчитана надежность результатов тестирования. В педагогической литературе тест считается надежным, если значение коэффициента альфа Кронбаха не менее 0,7. Во всех использованных вариантах значение коэффициента надежности находилось в пределах от 0,7 до 0,78, что свидетельствует о достаточной степени надежности проектируемых тестов.

Формирование базы шкалированных тестовых заданий по физике с использованием модели Раша послужило основой для разработки автоматизированной системы адаптивного тестирования. Автоматизированная система состоит из нескольких функциональных модулей: модуль формирования тестов, модуль сбора и обработки ответов учащихся, интерфейс администратора и интерфейс пользователя. В системе реализована ветвящаяся стратегия, при которой отбор заданий осуществляется непосредственно из базы с опорой на алгоритмы, прогнозирующие оптимальный уровень трудности последующего задания на основе результатов выполнения предыдущих. Такой подход обеспечивает индивидуализацию траектории тестирования и повышение точности оценки уровня подготовленности учащихся. Отличительной особенностью варьирующей ветвящейся стратегии адаптивного тестирования является пошаговая переоценка уровня подготовленности учащихся, осуществляемая после выполнения очередного тестового задания.

Варьирующая стратегия адаптивного тестирования включает три основных компонента, реализуемых последовательно на каждом этапе тестирования. Во-первых, осуществляется определение априорной оценки уровня подготовленности учащегося, во-вторых, производится отбор тестового задания, оптимального по своим параметрам с точки зрения соответствия текущему уровню подготовленности учащегося, и в-третьих, по результатам выполнения задания вычисляется апостериорная оценка уровня подготовленности, используемая для адаптации последующих шагов тестирования. Алгоритм, реализующий варьирующую стратегию, носит циклический характер и представлен на рисунке 4.

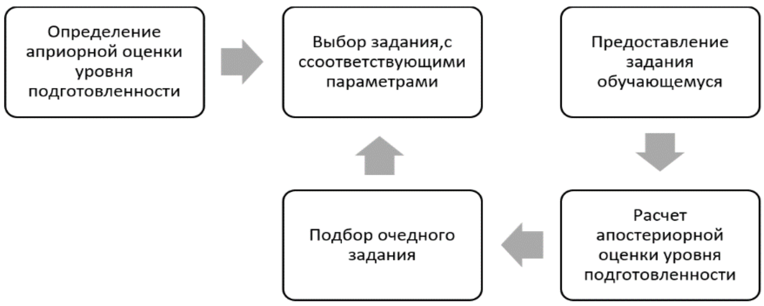

Рисунок 4 – Алгоритм варьирующего многошагового тестирования

Эффективность адаптивного тестирования, основанного на данной стратегии, определяется не только точностью подбора заданий по уровню их трудности, но и выбором априорной оценки уровня подготовленности учащегося, от которой зависит траектория теста на его начальных этапах. Стратегия адаптивного тестирования, основанная на модели Раша и априорной оценке уровня подготовленности, обеспечивает возможность индивидуализированного подбора заданий уже на начальных этапах формирования адаптивного теста. Выбор модели Раша обусловлен рядом методологических и прикладных факторов. В отличие от двух- (2PL) и трёхпараметрических (3PL) моделей теории IRT, модель Раша предполагает, что вероятность правильного ответа зависит *исключительно* от разности между уровнем подготовленности обучающегося и трудностью задания. При этом дифференцирующая способность заданий и вероятность угадывания рассматриваются как фиксированные параметры, что обеспечивает интерпретируемость шкалы, инвариантность оценок и упрощает процесс шкалирования заданий. В таблице 5 представлены ключевые различия между моделями по критериям интерпретируемости, сложности оценки параметров, вычислительных затрат и устойчивости к объему выборки.

Таблица 5 – Сравнение моделей IRT

Параметр	Модель Раша	2PL-модель	3PL-модель
Кол-во параметров на задание	1 (трудность задания)	2 (трудность, дифференцирующая способность)	3 (включая угадывание)
Сложность оценки	Низкая	Средняя	Высокая
Интерпретируемость	Высокая	Средняя	Ниже средней
Вычислительные затраты	Низкие	Средние	Высокие
Подходит для адаптивного тестирования	Да (простота реализации)	Да, но сложнее	Да, но ресурсоёмкая
Устойчивость при малых выборках	Высокая	Ниже	Низкая

Согласно данным, представленным в таблице, модель Раша требует меньших выборок для устойчивой оценки параметров, что особенно важно на начальных этапах разработки адаптивных тестов. Несмотря на ограничения по гибкости, её использование оправдано при построении базовой структуры адаптивного тестирования и позволяет обеспечить высокую надёжность измерений при соблюдении требований к *качеству заданий*.

Заключение

Исследование, проведенное с помощью программы Winsteps, позволило всесторонне оценить качество тестовых заданий. Полученные результаты подтверждают, что большинство заданий соответствуют предъявляемым требованиям и рекомендованы для использования в адаптивном тестировании. Выявленные случаи отклонений были тщательно проанализированы, что позволило внести соответствующие изменения и дополнения в задания для повышения качества инструмента.

Дополнительная экспертиза, проведенная по итогам апробации тестовых заданий, позволила выявить пробелы, которые коснулись не только тестологии, но и подготовленности учащихся. По итогам исследования было выявлено, что у большинства учащихся вызывали затруднения:

– задания по оптике, учащиеся не полностью понимают основные концепции волновой природы света;

– трудности с интерпретацией графиков: даже если учащиеся могут прочитать график, они испытывают трудности с интерпретацией его значений, не понимают взаимосвязи между переменными, не могут делать правильные выводы на основе данных.

По результатам апробации было проанализировано 10 вариантов тестов по физике. Определены параметры 300 тестовых заданий, из которых 278 рекомендованы для использования в адаптивном тестировании. 22 тестовых задания следует пересмотреть и провести повторную апробацию, так как данные задания (например, задание №29) вызвали затруднения у учащихся с высоким уровнем подготовленности. По итогам психометрического анализа 27 тестовых заданий (9%) были отредактированы, внесены незначительные изменения в условия заданий, дополнены графики и рисунки 3-х тестовых заданий, изменены дистракторы 15 тестовых заданий. 278 тестовых заданий в дальнейшем рекомендуется использовать в адаптивном тестировании с сохранением первоначального уровня трудности.

Результаты проведенного исследования могут быть положены в основу разработки цифровых систем адаптивного тестирования, интегрируемых в современные LMS (Learning Management Systems) и образовательные платформы. Алгоритмизированный подход к отбору и шкалированию заданий с применением модели Раша и инструмента Winsteps позволяет формировать банк тестовых заданий с объективной мерой трудности, пригодной для автоматического подбора заданий на основе уровня подготовленности обучающегося. Указанные инструменты обеспечивает совместимость систем управления базами данных с аналитическими системами, позволяя накапливать и обрабатывать большие массивы данных об обучении и оценке учебных достижений. В перспективе данная методология может быть расширена с использованием методов искусственного интеллекта для построения персонализированных траекторий обучения.

Список литературы

1. Liu Q. Survey of Computerized Adaptive Testing: A Machine Learning Perspective. [Электрон. ресурс]. – 2024. – URL: https://arxiv.org/pdf/2404.00712 (дата обращения 12.12.2024).
2. Developing of computerized adaptive testing to measure physics higher order thinking skills of senior high school students and its feasibility of use / E. Istiyono et al // European Journal of Educational Research. – P. 180-183: https://doi.org/10.12973/eu-jer.9.1.91.
3. Бондаренко С.В. Использование ИИ для построения адаптивных тестов в цифровых платформах / С.В. Бондаренко, С.В. Мальцева // Информационные технологии в образовании. – 2023. – № 4. – С. 12-21.
4. Аванесов В.С. Критерии качества педагогических измерений [Электрон. ресурс]. – 2021. URL*:* https://testolog.narod.ru/Theory75.html *(*дата обращения: 02.03.2025).
5. Бунимович А.Г. Информационные технологии адаптивного тестирования: алгоритмы и архитектура систем / А.Г. Бунимович, Н.С. Герасимова, И.Ю. Шевченко // Вестник НГУ. Серия: Информационные технологии. – 2021. – №19(3). – С. 50-65.
6. Аванесов В.С. Item response theory: основные понятия и положения, [Электрон. ресурс]. – 2020. – URL*:* https://testolog.narod.ru/Theory67.html (дата обращения: 13.03.2025 г.).
7. Челышкова М.Б. Теоретико-методологические и технологические основы адаптивного тестирования в образовании: дис. док. пед. наук: 13.00.01 / Челышкова Марина Борисовна. – Москва, 2001. – С. 324.
8. Аванесов В.С. Дистракторный анализ, [Электрон. ресурс]. – 2005 - *URL:* https://testolog.narod.ru/Theory46.html (дата обращения: 20.02.2025 г.).
9. Adegoke B.A. Comparison of Item Statistics of Physics Achievement Test using Classical Test and Item Response Theory Frameworks, Institute of Education / B.A. Adegoke // Journal of Education and Practice, University of Ibadan, Nigeria. – 2013. – № 22.
10. Нейман Ю.М. Введение в теорию моделирования и параметризации педагогических тестов / Ю.М. Нейман, В.А. Хлебников. – Москва, 2000. – 165 с.

References

1. Liu Q. Survey of Computerized Adaptive Testing: A Machine Learning Perspective. [Ehlektron. resurs]. – 2024. – URL: https://arxiv.org/pdf/2404.00712 (data obrashcheniya 12.12.2024). (In English).
2. Developing of computerized adaptive testing to measure physics higher order thinking skills of senior high school students and its feasibility of use / E. Istiyono et al // European Journal of Educational Research. – R. 180-183: https://doi.org/10.12973/eu-jer.9.1.91. (In English).
3. Bondarenko S.V. Ispol'zovanie II dlya postroeniya adaptivnykh testov v tsifrovykh platformakh / S.V. Bondarenko, S.V. Mal'tseva // Informatsionnye tekhnologii v obrazovanii. – 2023. – № 4. – S. 12-21. (In Russian).
4. Avanesov V.S. Kriterii kachestva pedagogicheskikh izmerenii [Ehlektron. resurs]. – 2021 - URL: https://testolog.narod.ru/Theory75.html (data obrashcheniya: 02.03.2025). (In Russian).
5. Bunimovich A.G. Informatsionnye tekhnologii adaptivnogo testirovaniya: algoritmy i arkhitektura sistem / A.G. Bunimovich, N.S. Gerasimova, I.YU. Shevchenko // Vestnik NGU. Seriya: Informatsionnye tekhnologii. – 2021. – №19(3). – S. 50-65. (In Russian).
6. Avanesov V.S. Item response theory: osnovnye ponyatiya i polozheniya, [Ehlektron. resurs]. – 2020 – URL: https://testolog.narod.ru/Theory67.html (data obrashcheniya: 13.03.2025 g.). (In Russian).
7. Chelyshkova M.B. Teoretiko-metodologicheskie i tekhnologicheskie osnovy adaptivnogo testirovaniya v obrazovanii: dis. dok. ped. nauk: 13.00.01 / Chelyshkova Marina Borisovna. – Moskva, 2001. – C. 324. (In Russian).
8. Avanesov V.S. Distraktornyi analiz, [Ehlektron. resurs]. – 2005 – URL: https://testolog.narod.ru/Theory46.html (data obrashcheniya: 20.02.2025 g.). (In Russian).
9. Adegoke B.A. Comparison of Item Statistics of Physics Achievement Test using Classical Test and Item Response Theory Frameworks, Institute of Education / B.A. Adegoke // Journal of Education and Practice, University of Ibadan, Nigeria. – 2013. – № 22. (In English).
10. Neiman YU.M. Vvedenie v teoriyu modelirovaniya i parametrizatsii pedagogicheskikh testov / YU.M. Neiman, V.A. Khlebnikov. – Moskva, 2000. – 165 s. (In Russian).

А.М. Искакова[1*], О.С. Салыкова[1], Н.А. Дидарбекова[2], А.А. Артыкбаева[1]
[1]Ахмет Байтұрсынұлы атындағы Қостанай өңірлік университеті,
110000 Қазақстан Республикасы, Қостанай қ, А. Байтұрсынов көшесі 47
[2]Қазақстан Республикасы Ғылым және жоғары білім министрлігі
«Ұлттық тестілеу орталығы»,
010011 Қазақстан Республикасы, Астана қ, Родниковый көшесі 1/1
*e-mail: n.a.almira.24@gmail.com

ФИЗИКА ПӘНІНЕН ТЕСТ ТАПСЫРМАЛАРЫН RASCH МОДЕЛІ НЕГІЗІНДЕ КОМПЬЮТЕРЛІК АДАПТИВТІК ТЕСТІЛЕУ ҮШІН ШКАЛАЛАУ

Адаптивті тестілеу – білімді цифрлық бағалаудың ең тиімді тәсілдерінің бірі болып табылады, себебі ол тапсырмаларды оқушының дайындық деңгейіне байланысты автоматты түрде іріктеу арқылы бағалау үдерісін жекешелендіруге мүмкіндік береді. Мұндай тестілеуді жүзеге асырудың негізгі элементтері: шкалаланған тапсырмалар банкі, адаптивтеу алгоритмі және арнайы бағдарламалық қамтамасыз ету. Сапалы тапсырмалар банкін қалыптастыру үшін оларды адаптивті жүйелерде қолдануға жарамдылығын бағалауға мүмкіндік беретін алдын ала психометриялық талдау қажет.

Бұл мақала физика пәні бойынша тест тапсырмалары жиынтығын Раш моделі негізінде эмпирикалық талдауға арналған. Зерттеу барысында тапсырмалар репрезентативті оқушылар іріктемесінде апробациядан өтіп, Winsteps бағдарламасы арқылы шкалалау жүргізілді. Әр тапсырма бойынша күрделілік параметрлері, модельмен сәйкестік көрсеткіштері және корреляциялық сипаттамалар анықталды. Адаптивті тестілеу талаптарына сәйкес келмейтін тапсырмалар анықталып, қорытынды банк құрамынан алынып тасталды. Нәтижесінде, статистикалық тұрғыдан тұрақты сипаттамаларға ие және білімді компьютерлік адаптивті жүйелерде бағалауға жарамды тапсырмалар жиынтығы қалыптастырылды.

Ұсынылған нәтижелер алынған тапсырмалар банкін ақпараттық білім беру жүйелері мен цифрлық платформаларға интеграциялау мүмкіндігін растайды. Алдағы жарияланымдарда нақты уақыт режимінде адаптивті тестілеу құру алгоритмдерін, сондай-ақ шкалаланған параметрлерді ескере отырып, тесттерді автоматты түрде құрастыратын бағдарламалық жасақтаманы әзірлеу жоспарлануда. Жүргізілген жұмыс оқу жетістіктерін бағалауға арналған тиімді цифрлық құралдарды жасауға негіз болады.

Түйін сөздер: *адаптивті тестілеу, Раш моделі, психометриялық талдау, тапсырмаларды шкалалау, білім берудегі ақпараттық технологиялар, Winsteps, білімді автоматтандырылған бағалау.*

A. Iskakova[1*], O. Salykova[1], N. Didarbekova [2], A. Artykbaeva[1]
[1]A. Baitursynov Kostanay Regional University,
110000 Republic of Kazakhstan, Kostanay, st. Baitursynov 47,
[2]National Testing Center of the Ministry of Science and
Higher Education of the Republic of Kazakhstan
010011 Republic of Kazakhstan, Astana, Rodnikovaya 1/1 st.
*e-mail: n.a.almira.24@gmail.com

SCALING PHYSICS TEST ITEMS FOR COMPUTERIZED ADAPTIVE TESTING BASED ON THE RASCH MODEL

Adaptive testing is one of the most effective approaches to digital knowledge assessment, providing personalization through the automated selection of test items tailored to the examinee's proficiency level. The key components of such testing include: a bank of scaled test items, an adaptation algorithm, and specialized software. Developing a high-quality item bank requires preliminary psychometric analysis to evaluate their suitability for use in adaptive systems.

This article presents an empirical analysis of a set of physics test items using the Rasch model. The study involved piloting the items on a representative sample of students, followed by scaling using the Winsteps software. For each item, difficulty parameters, model-fit indices, and correlation characteristics were determined. Items that did not meet the requirements of adaptive testing were identified and excluded from the final bank. As a result, a set of items with stable statistical properties was formed, suitable for further use in computerized adaptive knowledge assessment systems.

Сведения о документе

Тип документа	Внутренний документ
Номер и дата документа	15-10-4-2-01/4058 от 10.11.2025
Ссылка на документ	https://uto.workspace.kz/storage/document_attachments/f6Hojx81h35DQV7W0fNP5x74UGCYqeJLNDKbQyED.pdf
Отправитель	РГП на ПХВ "НАЦИОНАЛЬНЫЙ ЦЕНТР ТЕСТИРОВАНИЯ"
Автор	Каргалдаева М. Д., Руководитель (тел: , email: moldirk2018@gmail.com

Лист согласования

ФИО	Дата и время	Результат	ЭЦП
Алтыбаева Шугыла Болатовна	2025-11-10 14:53:40	Согласован	Нет

Лист подписания

Дидарбекова Наужан Абдисатаровна	2025-11-10 15:29:24	Подписан	Нет

Лист регистрации

Лист отправки

Лист корреспондентов